

# **Examining the Myth of Accountability, High-Stakes Testing, and the Achievement Gap**

Julian Vasquez Heilig  
California State University, Sacramento

T. Jameson Brewer  
University of North Georgia

Jimmy Ojeda Pedraza  
California State University, Sacramento

*Journal of Family Strengths*, 18(1), (2018), Article 9  
Available at: <https://digitalcommons.library.tmc.edu/jfs/vol18/iss1/9>

This research was supported by a grant from the Institute of Educational Sciences, Pathways to the Educational Sciences program, grant #R305B160016

## **Abstract**

This paper examines and critiques the accountability movement, high-stakes testing, and their relationship to the achievement gap. Analyzing the issues in the context of Texas, the paper discusses dropout rates that were incorrectly identified and reported, the role of courts and specific court cases in high-stakes testing, ethical considerations, social implications and social stratification, and debunks the myth of accountability as an equalizer. Authors conclude that the sorting at the root of high-stakes testing has neither closed the achievement gap nor fomented meaningful accountability or success.

### **Popularity and Common Belief: Birth of Texas-Style Accountability**

In the late 1990s, students of color in the large, urban high schools in Houston were reporting that they had 0% dropouts, and it was claimed that the achievement gap on the Texas Assessment of Academic Skills (TAAS) was closing rapidly. Education reformers attributed all of this purported success directly to Texas's implementation of high-stakes testing and accountability (Vasquez Heilig & Darling-Hammond, 2008). The Houston Independent School District and many other traditionally underperforming districts across the state were suddenly a success—it was a Texas miracle (Haney, 2000). But had Houston, and Texas, really experienced a miracle that would justify codifying high-stakes testing and accountability for every student in the entire nation?

Although the standards, testing, and accountability education reform movement is firmly situated as an offspring of the 1983 release of *A Nation at Risk* (ANAR), surely the passage of the No Child Left Behind (NCLB) Act of 2001 was rooted in policy making in Texas (Vasquez Heilig, Brewer, & White, 2018). In the 1980s and 1990s, there was a concerted push by Texas policymakers and business leaders to reform the state's schools (Vasquez Heilig & Darling-Hammond 2008). Texas was one of the earlier states to develop statewide testing systems during the 1980s, adopting minimum competency tests for school graduation in 1987 (Carnoy, Loeb, & Smith, 2003). In the early 1990s, the Texas Legislature passed Senate Bill 7 (1993), which mandated the creation of Texas-style public school accountability to rate school districts and evaluate campuses. Signed into law by Democratic Governor Ann Richards in 1993, S.B. 7 represented a bipartisan attempt to remedy the state's educational woes as it was passed by a wide margin in both the Texas House and Senate.

The first Texas accountability system, an information forum that used test scores and other measures of student progress to determine whether school districts should remain accredited by the state, was implemented in 1994. The Texas accountability system was undergirded by data in the Public Education Information Management System (PEIMS), a state-mandated curriculum, and statewide standardized testing to measure student proficiency in core subjects.

From 1995 to 1999, test-based accountability commenced in Texas under Governor George W. Bush. During this period, educational policy in the state evolved beyond implementing district-level consequences to applying a variety of sanctions on teachers, principals, and schools. The state also saw the promulgation of higher stakes for students, such as the abolition of social promotion, which is automatic grade progression. For

example, in Houston, Superintendent Rod Paige utilized TAAS and Stanford 9 test scores to determine whether students should advance to the next grade (Vasquez Heilig & Darling-Hammond, 2008).

The prevailing theory of action underlying Texas-style high-stakes testing and accountability ratings was that schools and students held accountable to these measures would automatically increase their educational output as educators tried harder, schools adopted more effective methods, and students learned more. Pressure to improve test scores would produce genuine gains in student achievement (Scheurich, Skrla, & Johnson, 2000). As test-based accountability commenced in Texas, achievement gains across grade levels conjoined with increases in high school graduation rates and decreases in dropout rates brought nationwide acclaim to the Texas accountability “miracle” (Haney, 2000).

The Texas miracle narrative was supported by high-stakes testing trends purportedly showing that African American and Latina/o students were closing the achievement gap on state-mandated tests over time. The first generation of Texas-style accountability relied on the TAAS from 1994 to 2002. For example, African Americans increased their achievement on the TAAS Exit Math; whereas only 32% met minimum standards in 1994, 85% did so by 2002. Concurrently, the percentage of Latinas/os meeting minimum standards increased from 40% to 88%. Although an achievement gap between minorities and whites remained, the gaps for Latinas/os and African Americans narrowed to 8% and 11%, respectively, between 1994 and 2002. Despite apparent success on the state-controlled TAAS test, large gains were not reflected in other national comparative exams, such as the National Assessment of Educational Progress (NAEP), American College Test (ACT), and Scholastic Achievement Test (SAT) (Vasquez Heilig, Jez, & Reddick, 2012).

### **Foundations and Literature: Accountability and High-Stakes Testing**

Early on, the research literature echoed the administrative progressive ideals that the long-term implications of accountability pointed to increased efficiency and achievement (Cohen, 1996; Smith & O'Day, 1991); however, others, positing Deweyan ideals, argued that testing would ultimately narrow the curriculum and negatively affect classroom pedagogy (McNeil & Valenzuela, 2001; Valencia & Bernal, 2000). Nevertheless, at the point of the national implementation of NCLB, the Texas miracle was the primary source of evidence, fueling the notion that accountability created more equitable schools and districts by positively affecting the long-term success of low-performing students (Nichols, Glass, & Berliner, 2006). In theory, accountability spurs high schools to

increase education output for all students, especially for African American and Latina/o students, who have been historically underserved by U.S. schools. Yet the question remains: Do policies that reward and sanction schools and students based on high-stakes test scores improve African American and Latina/o student outcomes over the long term?

We've already discussed testing before NCLB, so we now examine dropout data after the passage of NCLB to consider an additional measure of success. In 2005, when Texas began to use the National Center for Education Statistics (NCES) dropout definition for leaver reporting, the yearly count tripled for Latinas/os and quadrupled for African Americans (Vasquez Heilig et al., 2012). Clearly, Latinas/os and African Americans were overrepresented in the underreporting of yearly dropouts. In the 1998-1999 school year, the Texas Education Agency (TEA) introduced the tracking of individual students in cohorts between grades 9 and 12. African American and Latina/o cohort dropout rates halved between 1999 and 2005. However, after 2005, with use of the NCES dropout standard for leaver reporting, a 100% increase in the number of publicly reported dropouts occurred in Texas (Vasquez Heilig et al., 2012).

Notably, the cohort dropout rates more than doubled for African Americans and Latinas/os after adoption of the NCES standard. These numbers align with empirical research critical of the TEA's publicly reported dropout numbers (Losen, Orfield, & Balfanz, 2006; Vasquez Heilig & Darling-Hammond, 2008) and suggests that the number of students who left was underreported for quite some time by the state, especially when it came to African American and Latina/o populations. In summary, after NCLB, Texas did not experience an educational miracle, and the TEA vastly misrepresented the Lone Star State's success during the pre- and post-NCLB accountability eras.

### **Legal Implications: Accountability, High-Stakes Testing, and the Courts**

For about a hundred years, high-stakes standardized tests have been used to sort and track students in the United States. The use of tests was spurred early on by the racist eugenics movement to affirm its belief that one race was intellectually superior to another (Sacks, 1999). The first and most influential federal legal challenge in terms of high-stakes testing was *Debra P v. Turlington* (1981). The case was brought the National Association for the Advancement of Colored People (NAACP) on behalf of African American students who had failed the Florida high school exit exam. The NAACP argued in the lawsuit that students were not given enough notice and that the test was racially unfair. Furthermore, "at the

time of the 1979 hearing, after three test administrations, the failure rate of Black students was approximately 10 times greater than that of White students” (*Debra P. v. Turlington*, 1984, p. 1405). The court ruled in favor of the state but imposed two requirements on the schools: (1) Schools had to give students sufficient notice of the exam and (2) had to demonstrate that the subject matter that needed to be learned to pass the exam was in fact taught at the school. The court concluded that the “state may condition the receipt of a public-school diploma on the passing of a test so long as it is a fair test of that which was taught” (*Debra P. v. Turlington*, 1981, p. 406).

We now discuss two other notable challenges to high-stake testing in state courts. *Student No. 9 v. Board of Education* and *Valenzuela v. O’Connell* were two state court challenges that resulted in testing policy changes. In *Student No. 9 v. Board of Education*, the seniors graduating in the state of Massachusetts in 2003 challenged the state’s high school exit exam, known as the Massachusetts Comprehensive Assessment System (MCAS). The students argued that the MCAS “violated both due process and equal protection under the state constitution” (Holme & Vasquez Heilig, 2012). They believed that the test was unlawful and did not appropriately test their knowledge. The Massachusetts Supreme Judicial Court ruled against the students. However, the school was required to provide written notice to students if they failed the test, provide retesting opportunities, improve access and instruction for English language learners (ELLs) and disabled students, take specific action to reduce the number of dropouts, and reduce restrictions on appeals for students who fail (Massachusetts Department of Elementary and Secondary Education, 2006).

In *Valenzuela v. O’Connell*, California students challenged the California High School Exit Exam (CAHSEE). The students stated that the CAHSEE was unconstitutional because low-income and minority students were not given the same access to educational resources as their more affluent counterparts. The Alameda County Superior Court judge sided with the students. Ultimately, California Assembly Bill 347 passed, which required instruction services at no cost to students for those who had not passed the CAHSEE for two consecutive years after grade 12 (Holme & Vasquez Heilig, 2012). Furthermore, the bill required the local county office of education to verify whether or not the districts were complying with the provisions of the settlement (California Education Code Section 52380.7a). In 2017, California passed Assembly Bill 830, following the recent trend among states to abandon high-stakes exit exams.

### **Ethical Implications: Accountability, High-Stakes Testing, and Gaming**

For two decades, on the basis of the Texas miracle, policymakers and pundits argued that high-stakes testing was the answer to improving the educational system in the United States. It is now very rare to hear these arguments. Therefore, it is important to ask who is harmed the most by high-stakes testing? When test scores are tied to a school's access to funds, schools have acted rationally, but perhaps unethically, to game the test and the accountability system (Vasquez Heilig & Darling-Hammond, 2008). The process of gaming the system has caused many students, many of them of low socioeconomic status, to be pushed out of school—essentially making schools averse to at-risk students (Vasquez Heilig, Young, & Williams, 2012). Gaming responses have not only wrongfully placed students in courses that are not beneficial but also have led to the assignment of low-scoring students to special education so that their scores are not factored into school accountability ratings (Allington & McGill-Franzen, 1992; Figlio & Getzer, 2002). Moreover, research has found that schools encourage low-scoring students to leave school, transfer to general equivalency diploma (GED) programs, or drop out so that their scores will not affect a school's funding (Haney, 2000; Smith, 1986). Thus, it is clear that when high-stakes testing is connected to school funding, schools have found a way to game the system, at the expense of our most vulnerable students.

### **Social Implications: Accountability, High-Stakes Testing, and Stratification**

There are also important large-scale social implications of accountability and high-stakes testing that purposefully affect social stratification. A noble lie is a myth or untruth, told by the elites in society to maintain social harmony and advance an agenda of social engineering. Plato described the noble lie in *The Republic* via a fictional tale about society being divided into sections of silver, iron, brass, and gold. High-stakes exams and accountability have essentially functioned as a noble lie because these “reforms” have not fomented equity or social justice but instead have codified a sorting mechanism of stratification—gold, silver, brass, and iron—or, in the parlance of NCLB, “Far Below Basic,” “Below Basic,” “Basic,” “Proficient,” and “Advanced.”

NCLB politically framed tests and accountability as civil rights; however, it entailed a variety of deleterious social implications. First, testing proponents went too far and caused a widespread backlash by requiring too many exams. For example, Texas required students to pass

15 exams to graduate from high school. This overemphasis on testing in Texas and elsewhere led to a national movement to “opt out” of testing. Second, exit exam failure means that students cannot receive a high school diploma, which has had a disparately large effect on low-income students and students of color, who are less likely to pass standardized exams. The fact that a student has not received a high school diploma because of failure to pass exit exams ultimately affects his or her lifetime earnings. Third, test-driven “accountability” linked to education reform has led to mass firings of teachers—primarily persons of color—in cities such as Chicago and New Orleans. Fourth, under NCLB, if a school does not raise the scores of students fast enough, the school can be closed or turned over to private operators. Fifth, high-stakes exams and accountability have led to a slowdown in the growth of student success in the United States. Reardon, Greenberg, Kalogrides, Shores, & Valentino (2012) found that improvement in our NAEP scores was more rapid before the implementation of NCLB and determined that it will take 80 more years to close the achievement gap. Finally, NCLB and test-driven accountability paved the way for the current conversation about school choice and the private control and privatization of education. The test-driven accountability approach to education not only deprived communities of democratically controlled neighborhood schools, it failed to improve educational outcomes while empowering and increasing segregation via school choice (Vasquez Heilig, 2013). Clearly, the social implications of high-stakes tests suggest that they are a noble lie.

### **The Dangers: Accountability and High-Stakes Testing**

Dworkin and Tobe (2015) point out that accountability concerns focus primarily on trust (or the lack thereof). They outline that trust is either organic or contractual. In organic trust, individuals trust one another through social relationships. The converse of organic trust is contractual trust, in which the terms and conditions of contracts outline the parameters of expectations and provide the opportunity for recourse should that trust be broken. Dworkin and Tobe suggest that the rise of accountability by way of standardized testing in American schools represents a shift from organic trust to a more rigid understanding of the relationship of society to the teacher as one of contractual trust. The trust relationship between a society and its teachers was, as Dworkin and Tobe point out, initially one of organic trust, in which it was understood that the best interests of their students informed the daily practices of teachers. However, the rise of standardized testing as a mechanism for greater accountability represents not only a shift toward a contractual trust arrangement but also suggests

that teachers are primarily “motivated by self-interest at the expense of their students” (p. 184). The broader shift toward contractual trust and accountability in education coincides with the growth of the business ideology that has driven much of education reform nationally and internationally.

Again, with ideological roots in the hysteria trumpeted by the release of ANAR, a slew of policy prescriptions related to accountability began to focus even more on the nation’s schools, teachers, and students. The release of ANAR in the 1980s continued what had become an increasing distrust of teachers and schools following their apparent failure to allow us to beat the Soviets into space. The launch of Sputnik in the 1950s coincided with the rise of an accountability philosophy directed at governments, promoted by Milton Friedman, and ushered in a new era of pushing for more accountability (deMarrais, Brewer, Atkinson, Herron, & Lewis, in press). The release of ANAR renewed the fear that schools and teachers had failed our nation’s students—suggesting it would have been considered an act of war if another country had done to us what we had allowed our teachers and schools to do—because they were not being held accountable. In short, ANAR claimed that U.S. schools were trapped in mediocrity and were not necessarily operating efficiently or effectively. The passage of NCLB in the early 2000s—promoted by then President George W. Bush, who purportedly oversaw the “Texas miracle”—created a new era of high-stakes accountability directly linked to standardized testing.

The high-stakes testing accountability that came with NCLB and the incessant push to meet “adequate yearly progress” lest a school lose funding was followed by a rise in teach-to-the-test pedagogy. Additionally, many school districts in large urban centers found that the mandate to implement high-stakes testing was not accompanied by an increase in funds for targeting the out-of-school factors, like poverty, that inform student performance in school. As a result, educators in Atlanta, for example, were pushed or incentivized to change student answers on tests to avoid losing even more funding for the very schools that often received the least amount of funds.

However, the threat of losing funds is a necessary component of the push to inject market- and business-oriented ideology into schools. The rise of punitive measures after poor test results comes straight from the playbook of what educator Jesse Hagopian termed the “testocracy.” In a TEDx talk, Hagopian outlined the fundamental damage that the testing regime—or “testocracy”—does to students; the average student will take 112 standardized tests, many of which are high-stakes tests, between



kindergarten and the senior year of high school. The requirement to undergo this battery of exams results in students and teachers spending upward of 16 hours per week in test preparation or test taking (Hagopian, 2016).

Another dangerous component of high-stakes testing is the narrowing of curriculum, which is divided into atomized components geared specifically toward specific tests. The reductionistic practice of linking curriculum and testing puts constraints not only on teacher autonomy to direct and create curriculum but also on the time and flexibility needed to design a curriculum responsive to student interests. And while the reductionistic nature of testing and test preparation pedagogy likely encourages teacher burnout, as Dworkin and Tobe (2015) point out, the general shift toward contractual trust accountability in and of itself may also exacerbate teacher burnout.

High-stakes testing accountability is not limited to curriculum-specific testing. Increasingly, the average SAT score of students at a high school have become a metric for accountability across various levels. Yet, the SAT itself is mired in covert racial bias that traces its very roots back to the eugenics movement (Sacks, 1999) and the assumption that non-Whites are not as intelligent as Whites, regardless of their economic status (Hernstein & Murray, 1994).

### **Addressing and Debunking: Accountability and High-Stakes Testing**

In his discussion of the “testocracy,” Jesse Hagopian chronicles the rise of the opt-out movement that is growing across the country as educators and parents begin fighting back against the rise of standardized testing. In fact, the boycott of the Measures of Academic Progress (MAP) test that began in Hagopian’s high school came from a commitment to “refuse to do harm to students” (Hagopian, 2016).

Furthermore, much of the growth of “no excuses” charter schools and fast-entry teacher preparation programs like Teach For America has rested on the assumption that the best way to overcome poverty is to raise student test scores (Vasquez Heilig, Cole, & Springel, 2011). The logic, as it were, is that a student’s best opportunity to escape generational poverty is through schooling that reduces the process down to test scores. These assumptions intentionally overlook concepts in educational psychology (e.g., Maslow’s Hierarchy of Needs) and the effect that the pangs of poverty have on student performance in schools (Berliner & Biddle, 1995; Biddle, 2014; Brewer & Myers, 2015; Brill, 2011; Coleman, 1990; Coleman et al., 1966; Jencks et al., 1972; Ladson-Billings, 2006; Rothstein, 2004).

The assertion that the best way to alleviate poverty is to increase accountability by way of test scores (1) ignores the fact that two-thirds of all educational outcomes are informed by out-of-school factors (Rothstein, 2010) and (2) reduces poverty to individual failure. Operating under the myth of meritocracy, the assumption that test scores are the ticket out of poverty necessarily requires an assumption that the persistence of generational poverty is due not to systemic inequality but rather to bad teachers and a poor work ethic on the part of students—most often students of color. As a result, we must continue to push back against and debunk the detrimental myths surrounding the expansion of high-stakes testing. Doing so will require an ongoing discussion of the effects of out-of-school factors that testing simply does not address, in addition to further efforts by educators like Hagopian, who refuse to cause more harm to students by way of testing.

## **Conclusion**

In this article, we have outlined how notions of accountability and the achievement gap have relied upon the massive expansion of high-stakes exams in our nation's schools. The state of Texas has been a hotbed for experimentation with school reform, including the expansion of high-stakes testing. As explicated above, the "Texas miracle" never happened. Nevertheless, a decade of national education policy focused on high-stakes testing and accountability—despite the fact that the rise of high-stakes testing also involved considerable legal, ethical, and social considerations. Most importantly, Texas-style test and punish accountability manifested in various ways within schools and school culture across the nation via NCLB, which has undermined notions of trust within the teaching profession. The shift from organic to contractual trust has reimagined the role of the teacher to be that of a service provider who, being informed by his or her own self-interest, cannot be trusted to provide sufficient and quality education. The lack of trust that necessitates the need for contractual arrangements of accountability aligns with a business-oriented view of school reform and practices and pushes schools away from humanistic practices and toward market commodification.

Ideology dating back to the 1950s and Milton Friedman's assertion that government-run schools are innately inefficient and ineffective allowed reformers during the years and decades that followed to continue to find reasons to justify the implementation of policies of accountability. The logic behind the reductionistic nature of high-stakes testing is that it provides a standard quantified metric by which educators can, purportedly, gauge student improvement over time and compare them with one

another. And what follows from the ability to compare one student with another is the ability to compare one school with another, or one state with another. The goal of comparison is a key component of market-oriented notions of competition.

In conclusion, the practice of spending large amounts of time on test preparation and test taking must be reversed lest we continue on the path of maintaining schools solely as machinery for stratification. The foundation of high-stakes testing in the United States clearly has roots connecting the practice of sorting with the eugenics movement, which sought to “prove” through testing the existence of a racial hierarchy of intelligence. This foundation, in addition to market- and business-oriented ideology, has reinforced the racist under- and overtones of testocracy in the United States and has neither closed the achievement gap nor fomented meaningful accountability or success.

## References

- Allington, R. L., & McGill-Franzen, A. (1992). Unintended effects of educational reform in New York. *Educational Policy*, 6(4), 397–414.
- Berliner, D. C., & Biddle, B. J. (1995). *The manufactured crisis: Myths, fraud, and the attack on America's public schools*. Reading, MA: Addison-Wesley.
- Biddle, B. (2014). *The unacknowledged disaster: Youth poverty and educational failure in America*. Rotterdam, The Netherlands: Sense Publishers.
- Brewer, T. J., & Myers, P. S. (2015). How neoliberalism subverts equality and perpetuates poverty in our nation's schools. In S. N. Haymes, M. V. d. Haymes, & R. Miller (Eds.), *The Routledge handbook of poverty in the United States* (pp. 190–198). New York, NY: Routledge.
- Brill, S. (2011). *Class warfare: Inside the fight to fix America's schools*. New York, NY: Simon & Schuster.
- Carnoy, M., Loeb, S., & Smith, T. L. (2003). The impact of accountability policies in Texas high schools. In M. Carnoy, L. S. Siskin, & R. F. Elmore (Eds.), *The new accountability: High schools and high-stakes testing* (pp. 147–175). New York, NY: RoutledgeFalmer.
- Cohen, D. (1996). Standards-based school reform: Policy, practice, and performance. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in Education* (pp. 99-127). Washington, DC: Brookings Institution Press.
- Coleman, J. (1990). *Equality and achievement in education*. Boulder, CO: Westview Press.
- Coleman, J., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.
- Debra P. v. Turlington, 644 F.2d 397 (5th Cir., 1981).
- Debra P. v. Turlington, 730 F.2d 1405 (11th Cir., 1984).
- deMarrais, K., Brewer, T. J., Atkinson, J. C., Herron, B., & Lewis, J. (in press). *Philanthropy, hidden strategy, and collective resistance: A primer for concerned educators*. Borham, ME: Myers Education Press.
- Dworkin, A. G., & Tobe, P. E. (2015). Does school accountability alter teacher trust and promote burnout? In J. H. Ballantine & J. Z. Spade (Eds.), *School and society: A sociological approach* (5th ed.). Los Angeles, CA: Sage.

- Figlio, D. N., & Getzer, L. S. (2002). *Accountability, ability, and disability: Gaming the system?* Cambridge, MA: National Bureau of Economic Research.
- Hagopian, J. (2016). More than a score: Giving students a solid chance. Retrieved from <https://www.youtube.com/watch?v=gL64chNiuJQ>
- Haney, W. (2000). The myth of the Texas miracle in education. *Education Policy Analysis Archives*, 8, 41. doi:10.14507/epaa.v8n41.2000
- Hernstein, R., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York, NY: Free Press.
- Holme, J. J., & Vasquez Heilig, J. (2012). High-stakes decisions: The legal landscape of high school exit exams and the implications for schools and leaders. *Journal of School Leadership*, 22(6), 1177-1197.
- Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., . . . Michelson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York, NY: Basic Books.
- Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in U.S. schools. *Educational Researcher*, 35(7), 3–12.
- Losen, D., Orfield, G., & Balfanz, R. (2006). *Confronting the graduation rate crisis in Texas* (Civil Rights Project at Harvard University).
- Massachusetts Department of Elementary and Secondary Education. (2006). *Settlement allows MCAS to stand as graduation requirement* [Press release]. Retrieved from <http://www.doe.mass.edu/news/news.aspx?id=2907>
- McNeil, L., & Valenzuela, A. (2001). The harmful impact of the TAAS system of testing in Texas: Beneath the accountability rhetoric. In M. Kornhaber & G. Orfield (Eds.), *Raising standards or raising barriers? Inequality and high-stakes testing in public education* (pp. 127–150). Washington, DC: Brookings Institution Press.
- Nichols, S. L., Glass, G. V., & Berliner, D. C. (2006). High-stakes testing and student achievement: Does accountability pressure increase student learning?. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas*, 14.
- Plato. The Republic. (trans. 1942 by B. Jowett). In L. R. Loomis (Ed.), *The great dialogues* (pp. 217–477). Roslyn, NY: Walter J. Black.
- Reardon, S., Greenberg, E., Kalogrides, D., Shores, K., & Valentino, R. (2012). Left behind? The effect of no child left behind on academic achievement gaps (Working paper). Stanford, CA: Stanford University Press.

- Rothstein, R. (2004). *Class and schools: Using social, economic, and educational reform to close the Black-White achievement gap*. Washington, DC: Economic Policy Institute.
- Rothstein, R. (2010). Rothstein: Why teacher quality can't be only centerpiece of reform. Retrieved June 11, 2018 from <http://voices.washingtonpost.com/answer-sheet/school-turnaroundsreform/rothstein-on-the-manifestos-ma.html>
- Sacks, P. (1999). *Standardized minds: The high price of America's testing culture and what we can do to change it*. New York, NY: Da Capo Press.
- Skrla, L., Scheurich, J. J., & Johnson, J. F. (2000). *Equity-driven achievement-focused school districts: A report on systemic school success in four Texas school districts serving diverse student populations*. Austin, TX: Charles A. Dana Center.
- Smith, F. (1986). *High school admission and the improvement of schooling*. New York, NY: New York City Board of Education.
- Smith, M. S., & O'Day, J. (1991). Educational equality: 1966 and now. In D. A. Verstegen & J. G. Ward (Eds.), *Spheres of justice in education*. New York, NY: Harper Collins.
- Vasquez Heilig, J. (2013). Reframing the refrain: Choice as a Civil Rights issue. *Texas Education Review*, 1(1), 83–94.
- Vasquez Heilig, J., Cole, H., & Springel, M. (2011). Alternative certification and Teach For America: The search for high quality teachers. *Kansas Journal of Law and Public Policy*, 20(3), 388–412.
- Vasquez Heilig, J., & Darling-Hammond, L. (2008). Accountability Texas-style: The progress and learning of urban minority students in a high-stakes testing context. *Educational Evaluation and Policy Analysis*, 30(2), 75–110.
- Vasquez Heilig, J., Jez, S., & Reddick, R. (2012). *Is Texas leading its peers and the nation?: A decadal analysis of educational data*. Austin, TX: The Institute for Urban Policy Research and Analysis, University of Texas at Austin.
- Vasquez Heilig, J., Young, M., & Williams, A. (2012). At-risk student averse: Risk management and accountability. *Journal of Educational Administration*, 50(5), 562–585.
- Vasquez Heilig, J., Brewer, T. J., & White, T. (2018). What Instead?: Reframing the debate about charter schools, Teach For America, and testing. In R. Ahlquist, P. Gorski, & T. Montano (Eds.), *Assault on kids and teachers: Countering privatization, deficit ideologies*

*and standardization of U.S. schools* (pp. 201–217). New York, NY:  
Peter Lang.